Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON THE ACTUAL AND PROBABLE ERRORS OF INTERPOLATED VALUES DERIVED FROM NUMERICAL TABLES BY MEANS OF FIRST DIFFERENCES.

BY R. S. WOODWARD, C. E.

[Continued from page 149.]

Since

$$\int_{-(a+b)}^{-(a-b)} \frac{\varepsilon+a+b}{4ab}\, d\varepsilon = \frac{b}{2a},$$

$(a-b)$ will be the probable error or r when the relation between a and b is such that

$$\frac{b}{2a} = \tfrac{1}{4}, \text{ or when } b = \tfrac{1}{2}a.$$

In this case $r = a-b = a-\frac{1}{2}a = \frac{1}{2}a$. When $b > \frac{1}{2}a$ and $< a$, r will be given by (α); and when $b < \frac{1}{2}a$, r will be given by (β).

Evaluating (α) there results $r = \pm [a + b - \surd(2ab)]$, while (β) gives $r = \pm\frac{1}{2}a$. Therefore, when b lies between 0 and $\frac{1}{2}a$

$$r = \pm \tfrac{1}{2}a; \tag{11}$$

and when b lies between $\frac{1}{2}a$ and a

$$r = \pm [a + b - \surd(2ab)]. \tag{12}$$

In the problem of interpolation, $a = (1-t)\frac{1}{2}$ and $b = t\,.\,\frac{1}{2}$. Supposing $1-t > t$,

$$\frac{b}{2a} = \frac{t}{2(1-t)} = \tfrac{1}{4} \text{ for } t = \tfrac{1}{3}.$$

Therefore, according to equations (11) and (12),

$$\left.\begin{array}{llll} \text{when } t \text{ lies between } 0 & \text{and } \frac{1}{3}, & r = \pm \frac{1}{4}(1-t), \\ \text{“} \quad t \text{ “} \quad \text{“} \quad \frac{1}{3} & \text{“} \ \frac{2}{3}, & r = \pm \frac{1}{2}[1-\surd\overline{2t(1-t)}], \\ \text{“} \quad t \text{ “} \quad \text{“} \quad \frac{2}{3} & \text{“} \ 1, & r = \pm \frac{1}{4}t. \end{array}\right\} \tag{13}$$

It appears from these formulas that the probable error of an interpolated value of the first species is always less than the probable error of a tabular value. In other words, such interpolated values are more precise than tab. values. The following table gives the values of r corresponding to several values of t.

For	$t = 0$	or	1	$r = \pm 0.25$
“	$t = \frac{1}{10}$	“	$\frac{9}{10}$	$r = \pm 0.23$
“	$t = \frac{2}{10}$	“	$\frac{8}{10}$	$r = \pm 0.20$
“	$t = \frac{3}{10}$	“	$\frac{7}{10}$	$r = \pm 0.17$
“	$t = \frac{4}{10}$	“	$\frac{6}{10}$	$r = \pm 0.16$
“	$t = \frac{5}{10}$	“		$r = \pm 0.15$

As a statistical test of the accuracy of the laws of error expressed by eq's (8), (9), (10), 500 actual errors, $\varepsilon' = (1-t)\varepsilon_1 + t\varepsilon_2$, corresponding to interpolated values derived from a 5-place table of log's have been computed by means of a 7-place table. In each case t was taken equal to $\frac{1}{2}$, or the errors of interpolated mid-values were computed. The arguments used were the following numbers: 20005, 20035, 20065, 20105, 20135, 20165, 20205, etc., in the same order to 36635. Now for $t = \frac{1}{2}$, $a = b = \frac{1}{4}$, and hence by (8) and (10)

$\varphi(\varepsilon) = 4\varepsilon + 2$ for values of ε bet'n $-\frac{1}{2}$ and 0,
$\varphi(\varepsilon) = -4\varepsilon + 2$ " " " ε " $+\frac{1}{2}$ " 0.

From the integral $\int(2 \pm 4\varepsilon)d\varepsilon$, the percentages of the whole number of errors to be expected between the limits 0 and 0.1, 0.1 and 0.2, etc., up to 0.4 and 0.5 may be computed. The following table gives a comparison of these percentages with those derived from the 500 actual errors.

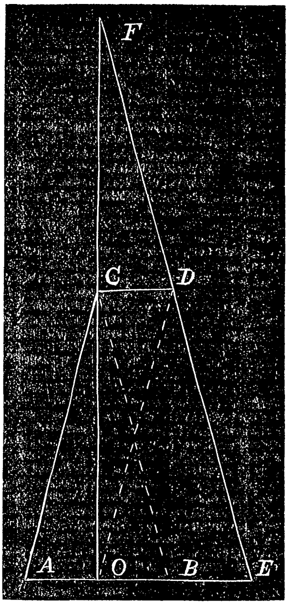
Between Limits.	Percent of whole number of errors.	
	By Theooy.	By Trial.
0.0 and 0.1	36	33.2
0.1 " 0.2	28	30.2
0.2 " 0.3	20	19.0
0.3 " 0.4	12	13.2
0.4 " 0.5	4	4.4
0.0 " 0.15	50	51.4

The average of the squares of all probable errors given by eq's (13) is

$$\left(\tfrac{1}{4}\right)^2 \left\{ 2\int_{2/3}^{1} t^2 dt + 4\int_{1/3}^{2/3} \left[1 - \sqrt{2t(1-t)}\right]^2 dt \right\} = 0.594\left(\tfrac{1}{4}\right)^2.$$

2. The consideration of the probable errors of interpolated values of the second species presents some peculiar features, which are, perhaps, more interesting than important. These features arise from the fact already stated that ε_3 in the formula $\varepsilon'' = (1-t)\varepsilon_1 + t\varepsilon_2 + \varepsilon_3$, is discontinuous for any value of t. Thus, when $t = \frac{1}{2}$, ε_3 must be either 0 or $\frac{1}{2}$ according as the tabulur difference is odd or even, and these two values of ε_3 are equally likely to occur. For $t = \frac{1}{2}$ therefore we may imagine any large number of resultant errors, ε'', to be made up of two distinct systems of error, namely, 1st that which embraces all cases in which $\varepsilon_3 = 0$, and 2nd that which embraces all cases in which $\varepsilon_3 = \frac{1}{2}$. In the first system the probable error is ± 0.15 as given by the second of (13) for $t = \frac{1}{2}$, since $\varepsilon_3 = 0$. In the second system all values of ε'' will be positive or negative according as ε_3 is taken $= +\frac{1}{2}$ or $= -\frac{1}{2}$; and the probable error will be $+\frac{1}{2}$ or $-\frac{1}{2}$ according as $\varepsilon_3 = +\frac{1}{2}$ or $-\frac{1}{2}$.

On the other hand, we may consider any large number of values of ε'', embracing both classes in which $\varepsilon_3 = 0$ and $\varepsilon_3 = \frac{1}{2}$, as one system. In such a system, supposing the value for ε_3 positive only, the actual errors will vary in magnitude between the limits $-\frac{1}{2}$ and $+1$. The law of facility for the errors in the system in which $\varepsilon_3 = 0$ only, is represented by the isosceles triangle *ABC*, the base *AB* being the axis of ε and the altitude *OC*, being the axis of $\varphi(\varepsilon)$. Likewise, the isosceles triangle *ODE* represents the law of facility for the system of errors in which $\varepsilon_3 = +\frac{1}{2}$ only, the origin being at *O* with axes *OE* and *OC* respectively. If the element-areas representing the probabilities of eq'l errors in the two systems be summed, the result will be the probability area corresponding to the combined system of errors. This area is represented by the trapezoid *ACDE*. Therefore



for values of ε between $-\frac{1}{2}$ and 0, $\varphi(\varepsilon) = 2 + 4\varepsilon$,
" " " ε " 0 " $+\frac{1}{2}$, $\varphi(\varepsilon) = 2$,
" " " ε " $+\frac{1}{2}$ " $+1$, $\varphi(\varepsilon) = 4 - 4\varepsilon$.

If we consider all possible errors in this system with reference to magnit'de only, the probability-curve becomes a straight line joining the points whose coordinates are (0, 4) and (1, 0), such line being represented by *EDF* in the above Figure. The law of facility is

$$\varphi(\varepsilon) = 4(1-\varepsilon).$$

To determine the probable error, r, we have the condition

$$\int_0^r 4(1-\varepsilon)d\varepsilon = \tfrac{1}{2}\int_0^1 4(1-\varepsilon)d\varepsilon = 1, \text{ whence}$$

$$4r - 2r^2 = 1 \text{ and}$$

$$r = 1 - \tfrac{1}{2}\sqrt{2} = +0.29.$$

In order to verify by an actual test this result, which indeed differs from the result reached by a distinguished mathematician, see § 5, the actual errors of 500 interpolated mid-values from a 5-place table of logarithms have been computed by means of a 7-place table. These mid-values correspond to the arguments mentioned under § 4, 1, above; and their actual errors correspond to ε'' in the formula $\varepsilon'' = (1-t)\varepsilon_1 + t\varepsilon_2 + \varepsilon_3$, for $t = \frac{1}{2}$. Now by the law of error just derived, one-half of these actual errors ought to be less than 0.29. Likewise, from the integral $\int 4(1-\varepsilon)d\varepsilon$, the percentages of the whole number of errors to be expected between the limits 0 and 0.2, 0.2 and 0.4 etc., to 0.8 and 1 can be computed. The following table affords a comparison of the theoretical and observed percentages.

Between Limits.	Per cent by Theo.	Per cent by Obs.
0 and 0.2	36	35.8
0 " 0.29	50	49.8
0.2 " 0.4	28	27.8
0.4 " 0.6	20	18.6
0.6 " 0.8	12	12.2
0.8 " 1.0	4	5.6

Again, suppose $t = \frac{1}{3}$. Then $\varepsilon_3 = 0, +\frac{1}{3}$ or $-\frac{1}{3}$, and these values of ε_3 are equally likely to occur. Any large number of actual errors, ε'', for $t = \frac{1}{3}$ may therefore be considered as a combination of three distinct systems corresponding respectively to the three different values of ε_3. For the system in which $\varepsilon_3 = 0$ only, $\varphi(\varepsilon)$ is given by equations (8), (9), (10) by putting $a = \frac{2}{3} \times \frac{1}{2} = \frac{2}{6}$, and $b = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$. For the system in which $\varepsilon_3 = +\frac{1}{3}$ only, $\varphi(\varepsilon)$ will result from (8), (9), (10) by giving to a and b the values just mentioned, and substituting for ε, $\varepsilon + \frac{1}{3}$; or in other words by moving the origin a distance $+ \frac{1}{3}$ to the left of its position for the system of errors in which $\varepsilon_3 = 0$ only. Likewise, for the system in which $\varepsilon_3 = -\frac{1}{3}$ only, $\varphi(\varepsilon)$ will result by substituting for ε in $\varphi(\varepsilon)$ for the first system, $\varepsilon - \frac{1}{3}$. See Fig., p. 149, $t = \frac{1}{3}$. If the areas representing the probabilities of equal errors in the three systems be summed, the probability-area for the original or combined system will result. This area will be a trapezoid whose base is $\frac{10}{6}$, whose side parallel to the base is $\frac{2}{6}$, and whose altitude is 3, the other two sides making angles with the base whose tangents are $+\frac{9}{2}$ and $-\frac{9}{2}$ respectively. Therefore

for errors lying between $-\frac{5}{6}$ and $-\frac{1}{6}$, $\varphi(\varepsilon) = \frac{1}{4}(15+18\varepsilon)$,
" " " " $-\frac{1}{6}$ " $+\frac{1}{6}$, $\varphi(\varepsilon) = 3$,
" " " " $+\frac{1}{6}$ " $+\frac{5}{6}$, $\varphi(\varepsilon) = \frac{1}{4}(15-18\varepsilon)$.

Hence the probable error will result from

$$\int_{-\frac{5}{6}}^{-r} \tfrac{1}{4}(15+18\varepsilon)d\varepsilon = \int_{+r}^{+\frac{5}{6}} \tfrac{1}{4}(15-18\varepsilon)d\varepsilon = \tfrac{3}{4},$$

which gives

$$r = \pm \tfrac{1}{6}[5 - \sqrt{(12)}] = \pm 0.26.$$

Similarly it may be shown that for $t = \frac{1}{4}$, in which case $\varepsilon_3 = 0, +\frac{1}{4}, \pm \frac{1}{2}$ or $-\frac{1}{4}$, the possible errors being considered with respect to mag. only,

$\varphi(\varepsilon) = \frac{1}{3}(24-16\varepsilon)$ for values of ε between 0 and $\frac{1}{4}$,
$\varphi(\varepsilon) = \frac{1}{3}(28-32\varepsilon)$ " " " ε " $\frac{1}{4}$ " $\frac{3}{4}$,
$\varphi(\varepsilon) = \frac{1}{3}(16-16\varepsilon)$ " " " ε " $\frac{3}{4}$ " 1,
and $r = \frac{1}{8}[7-\sqrt{(23)}] = 0.28.$

Likewise for $t = \frac{1}{5}$, in which case $\varepsilon_3 = 0, +\frac{1}{5}, +\frac{2}{5}, -\frac{2}{5}$ or $-\frac{1}{5}$, it may be shown that

$\varphi(\varepsilon) = \frac{25}{4}(\frac{9}{10}+\varepsilon)$ for values of ε between $-\frac{9}{10}$ and $-\frac{1}{10}$,
$\varphi(\varepsilon) = 5$ " " " ε " $-\frac{1}{10}$ " $+\frac{1}{10}$,
$\varphi(\varepsilon) = \frac{25}{4}(\frac{9}{10}-\varepsilon)$ " " " ε " $+\frac{1}{10}$ " $+\frac{9}{10}$,
and $r = \pm\frac{1}{10}[9-\sqrt{(40)}] = \pm 0.27.$

By extending the above process the probable errors in other and more complex cases might be determined. Without going into these details, however, it is evident that the probable error of an interpolated value of the second species will in general be greater than the probable error of a tabular value. For values of t such that $1 \div t$ is large relatively to the largest tabular difference in the table used, ε_3 may become insignificant in comparison with $(1-t)\varepsilon_1 + t\varepsilon_2$, and the probable error will be given in this case by formulas (13). At the limit, where $t = 0$, ε_3 vanishes, and $r = \pm 0.25$.

It is concluded, therefore, that the use of interpolated values of the sec'd species will not insure the highest precision possible with a given table. In precise computations, indeed, there can be no adequate reason for introducing the error ε_3 since the computer cannot avoid knowing its value.

3. By reference to equation (5) it appears that the probable error of an interpolated value of the third species is the same as that of a tablar value. Hence there must be on the whole a loss of precision in using interpolated values of this species.

4. For interpolated values of the fourth species the probable error will be given by equations (11) and (12) by making $a = \frac{1}{2}$ and assigning to b the limiting value of ε_4, which is subject to the condition mentioned in §3. Whatever this limiting value of ε_4 may be, the probable error of the interpolated value cannot be less than ± 0.25. When t is such that ε_4 may have any value between $+\frac{1}{2}$ and $-\frac{1}{2}$ the probable error will be

$$r = \pm(1-\tfrac{1}{2}\sqrt{2}) = \pm 0.29.$$

At the limit, where $t = 0$ and $\varepsilon_4 = 0$,

$$r = \pm 0.25.$$

HISTORICAL NOTE.

§ 5. Since at least three distinguished mathematicians have each considered one or more of the problems discussed in this paper, without having reached conclusions altogether accordant, a brief reference to their investigations may not be uninteresting.

1. Bessel, 1838, in Nos. 358 and 359 of the Astronomische Nachrichten, in an article entitled *Untersuchungen ueber die Wahrscheinlichkeit der Beobachtungsfehler*, has investigated the laws of error discussed in § 4, 1. This article is an admirable one and merits more attention than it seems to have received by writers on the theory of errors. In No. 529 of the Nachrichten, in a review of Captain Shortrede's 7-place tables of logarithms, Bessel has applied the laws of error defined by equations (8), (9), (10), to determine the probable error of an interpolated value of the first species. He observes that the proportional part may be computed by means of the tabular difference or by means of the true difference, giving rise to what we have called the first and third species. He then says

> Man bemerkt leicht, dass die *erste* methode die vortheilhaftere ist, aber ich erinnere mich nicht, dass der *wahrscheinliche* Fehler ihres Resultats methodische untersucht worden wære.

Bessel does not raise the question of the independence of the errors of consecutive tabular values. His analysis leads to equations (13). He does not consider the probable errors of interpolated values of the second species.

2. Bremiker, 1852, in the introduction to the Latin edition of his 6-place tables of logarithms has an important chapter entitled *De Erroribus, Quibus Computationes Logarithmicae Afficiuntur.* This chapt. aims to give a more comprehensive view of this subject than any work with which the writer is acquainted, and its omission in subsequent editions of Bremiker's tables would seem to have relegated it to unmerited obscurity. Among other problems Bremiker considers that of the probable errors of interpolated values of the second species. (See his § 12.) His treatment of this problem is, however, quite erroneous owing to the fact that he has considered ε_3 in $\varepsilon'' = (1-t)\varepsilon_1 + t\varepsilon_2 + \varepsilon_3$ continuous between the limits $+\frac{1}{2}$ and $-\frac{1}{2}$ for any value of t. He says, p. 69, his f_1, f_2, f_3 and ε being the same as our ε_1, ε_2, ε_3 and t, respectively,

> Sumatur, f_1 et f_2 esse errores logarithmorum tabulæ L_1 et L_2, ita intelligendos, ut $L_1 + f_1$ et $L_2 + f_2$ sint exacti logarithmorum valores, error logarithmi $L_1 + \varepsilon\Delta$ erit
>
> $$f_1 + \varepsilon(f_2 - f_1) + f_3$$
>
> si f_3 designat errorem, qui e producto decurtato $\varepsilon\Delta$ proficiscitur, f vero putatur quilibet valorum æque probabilium, qui inter $-\gamma(=\frac{1}{2})$ et $+\gamma$ jacent. . . . . Itaque primum quaeramus, quos valores summa
>
> $$(1-\varepsilon)f_1 + \varepsilon f_2 + f_3$$
>
> accipere possit, si pro illis f omnes valores, que esse possunt, aeque probabiles et intra $-\frac{1}{2}$ et $+\frac{1}{2}$ jacentes ponantur.

He then states that the limits of this sum are —1 and +1, and proceeds to determine the probable errors of interpolated values corresponding to several values of $\varepsilon = t$. His results are very different from those reached in §4, 2, inasmuch as they increase gradually from a minimum ± 0.26 for $t = \frac{1}{2}$, to a maximum ± 0.29 for $t = 0$. The second of these probable errors is curiously absurd, since the inference from it is that an interpolated value infinitely near a tabular value is less accurate than a tabular value.* It may be remarked that the probable error ±0.26 for $t = \frac{1}{2}$ does not correspond to the percentages of actual errors given in §4, 2.

3. F. G. W. Struve, 1860, in *Arc Du Méridian,* Tome I, p. 94, has considered the problem of the probable error of an interpolated value of the first species. He gives no analysis, but says briefly and somewhat obscurely

> L'erreur probable d'un logarithme donné immédiatement dans les tables est 0.25 de la septieme décimale; l'erreur probable d'un logarithme déduit par l' interpolation entre deux chiffres tabulaires n' est que $\sqrt{\frac{2}{3}}$ de l'autre erreur.

This statement may be interpreted thus. If in equation (2), ε_1 and ε_2 be supposed to conform to the ordinary law of error viz., $\varphi(\varepsilon)=(h\div\sqrt{a})e^{-h^2\varepsilon^2}$ the probable error corresponding to ε' will be

$$\tfrac{1}{4}[(1-t)^2+t^2]^{\frac{1}{2}} = \tfrac{1}{4}[1-2t+2t^2]^{\frac{1}{2}}.$$

Now the square root of the average of the squares of probable errors given by this formula is

$$\tfrac{1}{4}\left[\int_0^1(1-2t+2t^2)dt\right]^{\frac{1}{2}} = \tfrac{1}{4}\sqrt{\tfrac{2}{3}} = \pm 0.204.$$

Although ε_1 and ε_2 do not conform to the ordinary law of error, it is interesting to note that the above formula gives results which do not differ widely from the true results given by equations (13). Struve makes no reference to the writings of Bessel or Bremiker on this subject.

---

*[Because ε_3 must necessarily represent plus or minus 0, 1, 2, 3 or 4 units in the last place figure of the interpolated value

$$v' = (1-t)v_1 + tv_2,$$

and because any one of these digits is equally likely to occur, whatever the value of v' may be, the actual value of ε_3 is therefore independent of the *value* of v', depending only on the terminal digit in the numerical expression for v', and is therefore equally continuous with ε_1 and ε_2. But the probable error of the interpolated value will necessarily be increased by the addition of ε_3, and as its maximum occurs at the greatest distance of the interpolated value from mid-value, the absurdity above alluded to is not apparent.—Ed.]